National Aeronautics and Space Administration

GeneLab
Metadata & Processed data

Sam Gebre

GLDS Repository

GeneLab
Open Science for Exploration

Home | Repository | Data | Tools | Environmental Data | Submit Data | Help | Workspace

Search Data

All | GeneLab | NIH GEO | EBI PRIDE | ANL MG-RAST

Search Filters (GeneLab Only)

Project Type | Factors | Organisms | Assay Type | Clear

NASA GeneLab Data System

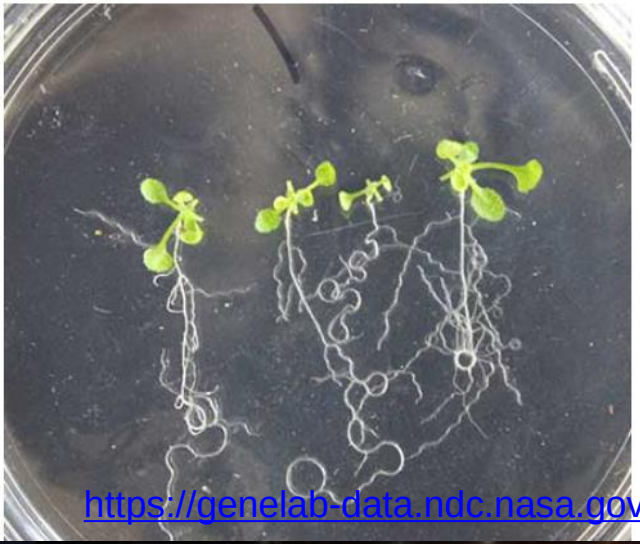
NASA GeneLab expands scientists' access to experiments onboard the International Space Station that explore the molecular response of terrestrial biology to spaceflight environments. Our mission is to maximize the utilization of the valuable biological research resources aboard the International Space Station by collecting genomic, transcriptomic, proteomic, and metabolomic data known as "omics".

Data Repository

The data repository hosts space biology and space-related datasets funded by multiple space agencies around the world. Use the search options above to filter through the GeneLab repository or to search across several databases.

Collaborative Workspace

GeneLab has customized a workspace for file sharing and access to data analysis tools. To access the workspace, users will need to create an account. Data analysis tool integration is ongoing and more tools will be available in future releases.



https://genelab-data.ndc.nasa.gov/genelab/

Searching the Repository

mouse liver transcriptomics x

All GeneLab

Search Filters (Gen

Project Type

Search results for: mou

Sort by Relevance

Rodent Research-3
https://genelab-data.ndc.

The Rod
Space to
weeks ol
sex and
Organism
Release/

STS-135: Mouse Liv
https://genelab-data.ndc.

Female
snap froz
ERCC co
Organism

STS-135 Liver Tran
https://genelab-data.ndc.

Mice wer
Samples
Organism

mouse liver transcriptomics x

All GeneLab NIH GEO EBI PRIDE ANL MG-RAST

Search results for: **mouse liver transcriptomics** using filter(s):

Total Search Results Found: **20**

Sort by Relevance 25

1

Rodent Research-3-CASIS: Mouse liver transcriptomic proteomic and epigenomic data

https://genelab-data.ndc.nasa.gov/genelab/accession/GLDS-137

The Rodent Research-3 (RR-3) mission was sponsored by the pharmaceutical company Eli Lilly and Co. and the Center for the Advancement of Science in Space to study the effectiveness of a potential countermeasure for the loss of muscle and bone mass that occurs during spaceflight. Twenty BALB/c 18-weeks old female mice (ten controls and ten treated) were flown to the ISS and housed in the Rodent Habitat for 39-42 days. Twenty mice of similar age sex and strain were used for ground controls housed ...

Organism: Mus musculus **Factor:** Spaceflight Treatment **Assay Type:** transcription profiling prot... **Accession:** GLDS-137 **PI/Contact:** Ruth Globus, NASA Gen...
Release/Publication Date: 28-Aug-2017

STS-135: Mouse Liver Transcriptomics using RNA-Seq

https://genelab-data.ndc.nasa.gov/genelab/accession/GLDS-173

Female C57BL/6 mice were flown onboard STS-135 for 13 days and returned to Earth for analysis. Livers were collected within 3-4 hours of landing and snap frozen in liquid nitrogen. Purified RNA samples that were used for microarray analysis for GLDS-25 were provided to GeneLab. GeneLab added ERCC control spike-in to the samples and performed RNA-Seq analysis.

Organism: Mus musculus **Factor:** Space Flight Absorbed R... **Assay Type:** transcription profiling **Accession:** GLDS-173 **PI/Contact:** Jonathan Galazka, NAS...
Release/Publication Date: 04-May-2018

Transcriptomics of single-cell and bulk sorted and micro-dissected mouse bone marrow and fetal liver cells

http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE89379

This SuperSeries is composed of the SubSeries listed below.

Organism: Mus musculus **Accession:** GSE89379 **PI/Contact:** Jean-Charles Boisset **Release/Publication Date:** 01-Nov-2016

mouse liver SWATH-MS

http://www.ebi.ac.uk/pride/archive/projects/PXD003266

Here, we have profiled 386 individuals in 80 cohorts of the BXD mouse genetic reference population across two environmental states through a metabolic phenotyping program including glucose response, exercise capacity, and cold response. To understand how the observed phenotypic differences are related to genetic variance, we generated a multilayered set of molecular phenotypes—genomics, transcriptomics, proteomics, and metabolomics across all cohorts, then modeled these molecular patterns with t...

Organism: Mus musculus (Mouse) **Accession:** PXD003266 **PI/Contact:** Yibo Wu **Release/Publication Date:** 13-Jun-2016

Search Data x

All GeneLab NIH GEO EBI PRIDE ANL MG-RAST

Search Filters (GeneLab Only)

Project Type | Factors | Organisms | Assay Type | Clear

Search results for: using filter(s): **[('Project Type = Flight Study' OR 'Project Type = Flight' OR 'Project Type = Spaceflight Project' OR 'Project Type = Spaceflight Study' OR 'Project Type = Spaceflight') AND ('Study Assay Measurement Type = transcription profiling') AND ('organism = Mus musculus')]**

Sort by Relevance 25

Total Search Results Found: **29**

1 2 Next >>

Spaceflight effects on the mouse retina: Histological gene expression and epigenetic changes after flight on STS-135

https://genelab-data.ndc.nasa.gov/genelab/accession/GLDS-87

We report the findings of an animal experiment onboard STS-135 investigating the molecular aspects of the impact of spaceflight on retinal biology by performing differential gene expression profiling between mice flown onboard STS-135 and their ground control counterparts.

Organism: Mus musculus **Factor:** Spaceflight Absorbed Ra… **Assay Type:** transcription profiling **Accession:** GLDS-87 **PI/Contact:** Susana Zanello
Release/Publication Date: 16-Dec-2016

Global gene expression analysis highlights microgravity sensitive key genes in soleus and EDL of 30 days space flown mice

https://genelab-data.ndc.nasa.gov/genelab/accession/GLDS-111

Microgravity exposure as well as chronic muscle disuse are two of the main causes of physiological adaptive skeletal muscle atrophy in humans and murine animals in physiological condition. The aim of this study was to investigate at both morphological and global gene expression level skeletal muscle adaptation to microgravity in mouse soleus and extensor digitorum longus (EDL). Adult male mice C57BL/N6 were flown aboard the BION-M1 biosatellite for 30 days on orbit (BF) or housed in a replicate ...

Organism: Mus musculus **Factor:** Microgravity Absorbed R… **Assay Type:** transcription profiling **Accession:** GLDS-111 **PI/Contact:** Dieter Blottner
Release/Publication Date: 15-Jan-2017

Biological and Metabolic Response in STS-135 Space-flown Mouse Skin

https://genelab-data.ndc.nasa.gov/genelab/accession/GLDS-116

Changes in gene expression profiles implicated in oxidative stress and in ECM remodeling in mouse skin were examined after space flight. The metabolic effects of space flight in skin tissues were also characterized.

Organism: Mus musculus **Factor:** Microgravity Absorbed R… **Assay Type:** metabolite profiling trans… **Accession:** GLDS-116 **PI/Contact:** Xiao Mao
Release/Publication Date: 28-Apr-2017

Data Repository

GLDS-194: Rodent Research-3-CASIS: Mouse retina transcriptomic data

GLDS-194: Rodent Research-3-CASIS: Mouse retina transcriptomic data

Source Accession Number
Total Data Volume: 599.6 GB

Submitted Date: 14-May-2018
Release Date: 06-Jun-2018

DESCRIPTION | PROTOCOLS | SAMPLES | ASSAYS/MEASUREMENTS | PUBLICATIONS | STUDY FILES

To view files, click on the folder of interest.

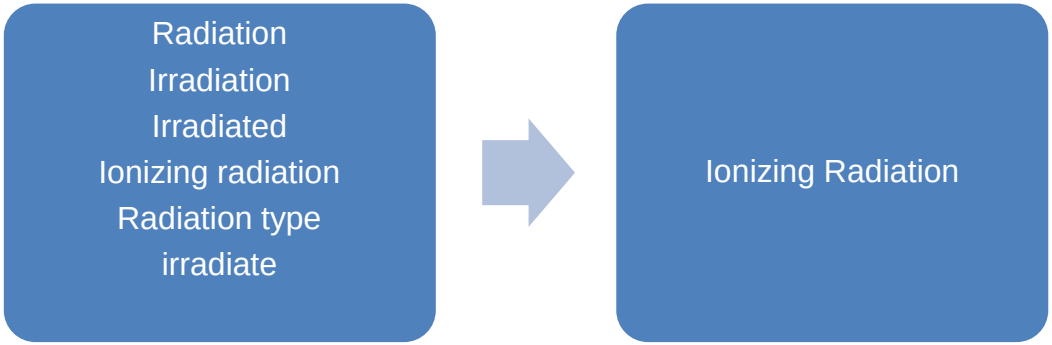
- All Files
  - Study Metadata Files
  - RNA-Seq Data

FILES	FILE SIZE	RESOURCE CATEGORY	RESOURCE NAME	RESOURCE DESCRIPTION
GLDS-194_metadata_RR3-RTN-ISA.zip	6.87 KB	metadata	Study Metadata Files	ISA-Tab format file. Tab delimited Investigation, Study, and Assay level metadata. See ISA-tools.org for format details and software.
GLDS-194_transcriptomics_RR3_BSL_B7_LRTN.tar	47.18 GB	transcriptomics	RNA-Seq Data	Raw data files in FASTQ format associated with this study.
GLDS-194_transcriptomics_RR3_BSL_B8_RRTN.tar	46.64 GB	transcriptomics	RNA-Seq Data	Raw data files in FASTQ format associated with this study.
GLDS-194_transcriptomics_RR3_BSL_B9_RRTN.tar	42.86 GB	transcriptomics	RNA-Seq Data	Raw data files in FASTQ format associated with this study.
GLDS-194_transcriptomics_RR3_BSL_B10_RRTN.tar	47.67 GB	transcriptomics	RNA-Seq Data	Raw data files in FASTQ format associated with this study.
GLDS-194_transcriptomics_RR3_FLT_F6_LRTN.tar	46.67 GB	transcriptomics	RNA-Seq Data	Raw data files in FASTQ format associated with this study.

Upcoming updates to the Repository

- Processed data will be added to the repository
  - Under Study Files

- Metadata normalization
  - Using ontologies, selected common terminology for fields by SMEs and Scientific community
    - Example: Study factor field for Ionizing Radiation



  - Standard Sample and file naming convention
    - GSM1287094 extract 1 → Dmel_OR_wo_FLT_uninfd_Rep1

Study Tab Files - Org

GLDS-1: Expression data from drosophila melanogaster

Source Accession Number E-GEOD-53196
Total Data Volume: 89.8 MB

Submitted Date: 10-Dec-2013
Release Date: 11-Dec-2013

DESCRIPTION | PROTOCOLS | SAMPLES | ASSAYS/MEASUREMENTS | PUBLICATIONS | STUDY FILES

To view files, click on the folder of interest.

- All Files
  - Microarray Data Files
  - Study Metadata Files
  - GeneLab Processed Microarray Data Files
    - Raw Data (Level 1)
    - Normalized Data (Level 1)
    - Differential Expression Analyses (Level 1)
    - Normalized Data (Level 1)/QC Reports (Level 2)

FILES	FILE SIZE	RESOURCE CATEGORY	RESOURCE NAME	RESOURCE DESCRIPTION
GLDS-1_array_Dmel_OR_wo_FLT_infdw-Bbas_Rep2_GSM1287107_raw.txt	353.2 KB	array	GeneLab Processed Microarray Data Files	Data files processed by the GeneLab Team, providing users data files processed using standardized methods.
GLDS-1_array_Dmel_OR_wo_FLT_infdw-Bbas_Rep3_GSM1287108_raw.txt	353.2 KB	array	GeneLab Processed Microarray Data Files	Data files processed by the GeneLab Team, providing users data files processed using standardized methods.
GLDS-1_array_Dmel_OR_wo_FLT_infdw-Bbas_Rep2_GSM1287107_raw.CEL	5.1 MB	array	GeneLab Processed Microarray Data Files	Data files processed by the GeneLab Team, providing users data files processed using standardized methods.

Microarray data files - Org

- **Raw files**
  - .CEL and .txt files
  - QC Reports
    - images.png
    - rawBoxplot.png
    - rawDensityDistributions.png
    - rawPCA.png
- **Normalized files**
  - normalized-annotated.rda
  - normalized-annotated.txt
  - normalized.txt
  - QC Reports
    - normBoxplot.png
    - normDensityDistributions.png
    - normPCA.png
    - normPlotMA.png
- **DGE files**
  - (Group 1)_VS_(Group 2).csv –- multiple files

Example of Processed Data - Normalized

ID	Dmel_OR_wo_FLT_uninfd_Rep1	Dmel_OR_wo_FLT_uninfd_Rep2	Dmel_OR_wo_FLT_uninfd_Rep3	Dmel_OR_wo_GC_uninfd_Rep1	Dmel_OR_wo_GC_uninfd_Rep2
NM_001273182	12.73436529	13.53763701	13.79472328	12.84755704	13.76895967
NM_001299640	13.07767325	13.28848484	13.15454729	12.5867939	13.9368497
NM_137584	8.853067596	8.933938761	8.600178682	8.663969717	8.595634732
NM_001259783	7.597094938	7.950533398	7.448852588	7.975693508	8.085393803
NM_001300537	7.301466824	7.902639581	7.512765582	7.84391638	7.97287487
NM_078946	4.447550387	4.32923845	4.46642468	4.444770459	4.418564324
NM_001104169	3.476285938	3.734146411	3.520484828	3.619259671	3.563646699
NM_137914	3.333274713	3.284510255	3.277965841	3.245932262	3.153632943
NM_137674	6.831546108	7.090095989	7.445701205	7.087853637	6.988525031
NM_078781	7.680847363	8.024377162	7.387456139	7.17879016	7.681871338
NM_001272532	8.865337111	8.858911864	8.539347437	8.50295581	8.599470174
NM_001144443	5.224297656	5.209867473	5.05337889	5.19471488	5.309226412
NM_001272886	6.645021556	6.690862075	8.534347219	6.351261382	6.597088019
NM_001201735	8.065852204	8.834707672	8.914475691	8.283007073	8.783242295
NM_001298306	4.807093079	4.476628524	5.016281316	4.631948502	4.419071171
NM_001104008	10.63333605	10.64024131	10.56589661	10.28910613	10.56874988
NM_141189	5.004836865	4.4272168	5.255189615	4.491323024	4.646059069
NM_001276161	8.973933258	9.821802949	10.48597845	9.033612817	9.498168257
NM_001300058	10.21745598	10.79485519	10.7250188	10.57363358	10.67663851
NM_001043265	3.943996394	4.617138807	4.79499779	4.039129563	4.026826331
NM_058056	4.928006366	5.788939768	6.656817364	5.811414346	5.3250182
NM_137472	6.724201192	7.006815195	6.253806681	6.664482884	6.847096702
NM_001201646	7.689893801	7.365592837	8.783415159	7.06976229	7.547382614
NM_001273491	5.664727266	5.676783611	5.57427777	5.375187187	5.331218967
NM_136552	5.029603944	4.862396826	4.604743017	4.790451419	5.411601235
NM_001275768	9.01056798	9.282647213	9.775363137	8.42458835	8.759034556
NM_143401	8.351893239	8.635076944	8.588446067	8.06583888	8.186445861

Annotated normalized expression data

Example of Processed data - DGE

(Flight).(Infected with B. bass)__VS__(Flight).(Infected with E. coli)

	logFC	AveExpr	t	P.Value	adj.P.Val	B
NM_143045	1.224182413	7.405480749	5.04530002	0.000116789	0.999786202	-3.717225032
NM_168735	1.147682177	6.810608914	4.72225404	0.000225654	0.999786202	-3.770742412
NR_002020	1.976283467	10.82795235	4.54440877	0.000325754	0.999786202	-3.802189031
NM_001258782	-0.994548074	7.185585288	-4.27178683	0.000574955	0.999786202	-3.853210646
NM_001297989	-0.593749515	3.989402653	-4.23885112	0.000616041	0.999786202	-3.859607965
NM_135852	0.98937063	6.853819288	4.10293731	0.000819654	0.999786202	-3.88654144
NM_001258983	-0.929340338	8.489856619	-3.82419928	0.001476572	0.999786202	-3.944452342
NM_001274605	1.005405446	5.607141129	3.81935171	0.001491802	0.999786202	-3.945490916
NM_140762	-0.722088645	7.961246103	-3.555013	0.002612025	0.999786202	-4.003701998
NM_079363	-0.923308515	8.743015415	-3.47661019	0.003084162	0.999786202	-4.021544983
NM_001275930	-1.010143075	7.724604606	-3.4758753	0.003088967	0.999786202	-4.021713432
NM_132102	-0.485949793	4.819642842	-3.4539197	0.003236035	0.999786202	-4.026756209
NM_132334	-0.778776721	5.878586958	-3.44705073	0.003283462	0.999786202	-4.0283379
NM_001104070	0.603645816	9.211636632	3.43531255	0.003366114	0.999786202	-4.031045212
NM_079406	-0.701633876	9.255580406	-3.42330912	0.003452775	0.999786202	-4.033819431
NM_001300461	1.127612754	9.9084038	3.38950222	0.00370895	0.999786202	-4.041663669
NM_001275502	0.499879585	7.215729991	3.37416981	0.003831271	0.999786202	-4.045236107
NM_001273214	1.109144474	10.49559077	3.35311032	0.004005834	0.999786202	-4.050157868
NM_001014696	-0.986219099	6.679948004	-3.34248972	0.004096838	0.999786202	-4.052646476
NM_134824	0.587185978	6.083973603	3.3155744	0.004336747	0.999786202	-4.058972529
NM_001298748	-0.632904094	9.893772781	-3.29509087	0.004528604	0.999786202	-4.063805203
NM_001297870	-0.615070739	6.354460299	-3.28533028	0.004622958	0.999786202	-4.066113523
NM_001298411	-0.626127077	5.806878173	-3.2569409	0.004908586	0.999786202	-4.072847375

Output Files from DGE

- (Ground Control).(Uninfected)__VS__(Ground Control).(Infected with E. coli).csv
- (Ground Control).(Uninfected)__VS__(Ground Control).(Infected with B. bass).csv
- (Ground Control).(Uninfected)__VS__(Flight).(Uninfected).csv
- (Ground Control).(Uninfected)__VS__(Flight).(Infected with E. coli).csv
- (Ground Control).(Uninfected)__VS__(Flight).(Infected with B. bass).csv
- (Ground Control).(Infected with E. coli)__VS__(Ground Control).(Uninfected).csv
- (Ground Control).(Infected with E. coli)__VS__(Ground Control).(Infected with B. bass).csv
- (Ground Control).(Infected with E. coli)__VS__(Flight).(Uninfected).csv
- (Ground Control).(Infected with E. coli)__VS__(Flight).(Infected with E. coli).csv
- (Ground Control).(Infected with E. coli)__VS__(Flight).(Infected with B. bass).csv
- (Ground Control).(Infected with B. bass)__VS__(Ground Control).(Uninfected).csv
- (Ground Control).(Infected with B. bass)__VS__(Ground Control).(Infected with E. coli).csv
- (Ground Control).(Infected with B. bass)__VS__(Flight).(Uninfected).csv
- (Ground Control).(Infected with B. bass)__VS__(Flight).(Infected with E. coli).csv
- (Ground Control).(Infected with B. bass)__VS__(Flight).(Infected with B. bass).csv
- (Flight).(Uninfected)__VS__(Ground Control).(Uninfected).csv
- (Flight).(Uninfected)__VS__(Ground Control).(Infected with E. coli).csv
- (Flight).(Uninfected)__VS__(Ground Control).(Infected with B. bass).csv
- (Flight).(Uninfected)__VS__(Flight).(Infected with E. coli).csv
- (Flight).(Uninfected)__VS__(Flight).(Infected with B. bass).csv
- (Flight).(Infected with E. coli)__VS__(Ground Control).(Uninfected).csv
- (Flight).(Infected with E. coli)__VS__(Ground Control).(Infected with E. coli).csv

Questions?

How to get the data?

- GeneLab API
- API wrapper - https://github.com/LankyCyril/genefab
  - Will be presented by Kirill Grigorev

Any requirements needed for your tool?

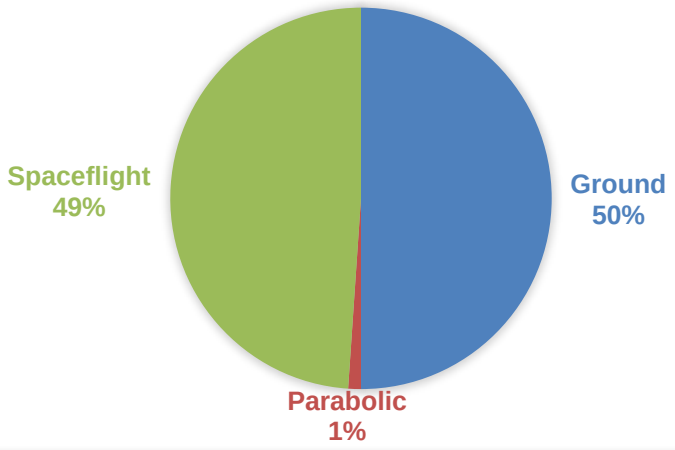
- Metadata
- Types of data
- DGE output data – correct format, fields,
  - Archive files
  - Group files
  - List all files (one by one)

Day 2: 9:00am – 10:30am

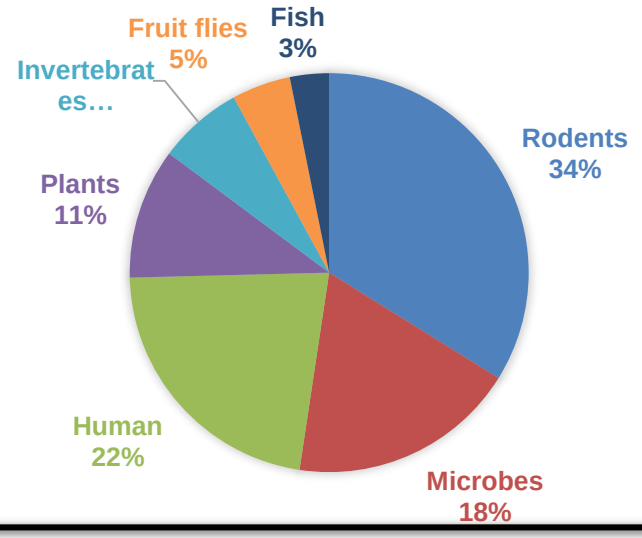
Interactive discussion on how to integrate the API with Basic Visualization tools

Overview: Database content

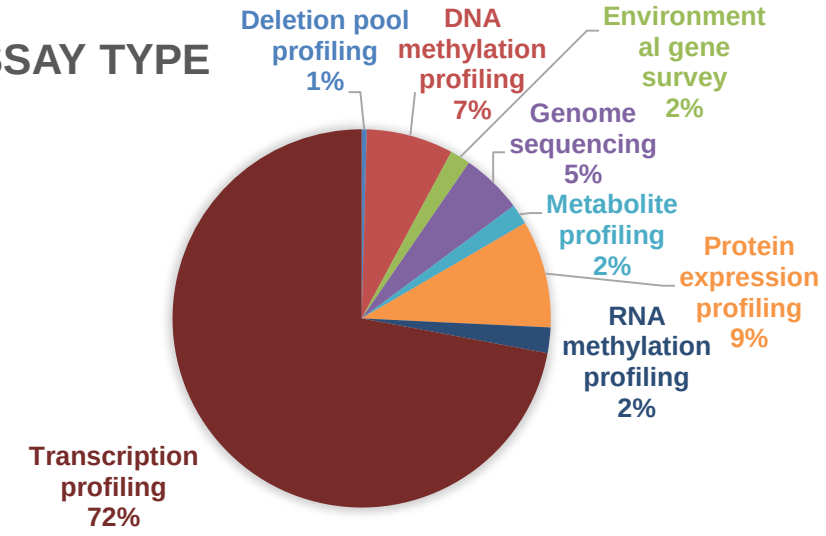
STUDY TYPE



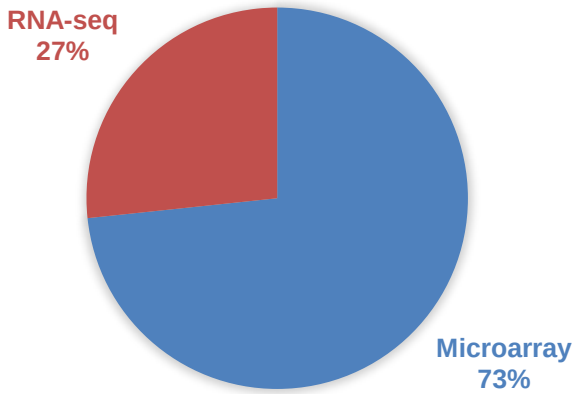
ORGANISM



ASSAY TYPE



TRANSCRIPTION PROFILING



Total # of studies: 189

Processed Data Schedule

- Microarray
  - Raw → Normalized expression levels →DGE
  - **Publish starting 5/22**
- RNA-Seq
  - Raw →Trimmed → Aligned → Normalized counts →DGE
  - **Publish starting 6/15**
- Metagenomics
  - Pipeline currently being defined/baseline
  - **Publish starting 8/1**
- Proteomics/Metabolomics
  - TBD
- Epigenomics
  - TBD